Tekst online vertalen

Traduttore, traditore?

Online vertaalmachines moeten de communicatie met de rest van de wereld vergemakkelijken. Met enkele muisklikken kan je woorden, zinnen, hele lappen tekst en zelfs complete webpagina's laten vertalen. Wat mogen we hiervan verwachten? Dicht bij elkaar blijven, en volg de gids! DIRK SCHOOFS



WAT DOEN WE?
- TALEN HERKENNEN EN TEKST VERTALEN

WAARMEE?
- EEN INTERNETVERBINDING

DUUR?
- 10 MINUTEN

MOEILIJKHEID?



Iedereen die ooit een online vertaler heeft geprobeerd, weet dat er op dit moment geen enkel systeem bestaat dat ook maar bij benadering even goed is als een menselijke vertaler. Hoogstwaarschijnlijk zal de situatie over 10 jaar niet anders zijn. Vertalen is immers veel meer dan woordjes omzetten. De huidige softwarematige vertalers kan je vergelijken met een middelbare scholier die zonder inzicht vanuit z'n eigen taal vertaalt naar het buitenlands.
Is elektronisch vertalen dan nutteloos? Neen. Die digitale vertaalmachines worden alsmaar knapper, en je kan ze wel gebruiken om mits enige goodwill te begrijpen wat de andere bedoelt. We nemen de proef op de som. Eerst proberen we of de zoekmachines de vreemde taal herkennen die we ingeven. Daarna trachten we enkele Russische paragrafen te begrijpen waar we op dit ogenblik geen jota van snappen. Ten derde proberen een Nederlandse tekst om te zetten naar vlot Engels. En tenslotte maken we gebruik van een service die een complete website vertaalt.

Scenario 1: Welke taal?

Om te achterhalen om welke taal het gaat, graven we bij **Xerox** naar de **Language Identifier** www.xrce.xerox.com/competencies/content-analysis/tools/guesser-ISO-8859-1.en.html. In het zoekvak plakken we achtereenvolgens drie regels tekst. We kopiëren de drie paragrafen telkens uit een buitenlandse online krant. De eerste is Lets, de tweede Kroatisch en de derde Maltees (zie afbeelding 1a). Telkens we bij **Xerox** op de knop **Guess Language** klikken, is het bingo: "Latvian, Croation, Maltese" Ongelofelijk.

1. Novembra sākumā Ankarā politiķiem, diplomātiem, politikas ekspertiem un žurnālistiem, populārākā frāze, ko pieminēja vai ik otrais runātājs, bija «vilciena katastrofa».

2. Drama gladi bila je tema današnjeg nedjeljnog Papinog nagovora vjernicima na Trgu Svetog Petra. Papa je oštro kritizirao postupke raznih vlada koje ostaju ravnodušne pred tom svjetskom tragedijom.

3. Indhil sfaccat anke fil-Kamra ta' l-Ahbarijiet tal-PBS GURNALISTI li jahdmu fil-kamra ta' l-ahbarijiet tal-Public Broadcasting Services (PBS), kif ukoll ohrajn li jahdmu ma' stazzjonijiet tar-radju privati, huma urtati ghall-ahhar minhabba pressjoni li qieghda ssir fuqhom mid-Dar Centrali tal-Partit Nazzjonalista.

1a

Welke talen worden hier gebruikt?

Evaluating your input...

Language	Match	Blue bars
Latvian	62.38%	
English	53.73%	
Latin	52.8%	
Italian	52.65%	
Swedish	50.17%	
French	47.98%	
Finnish	47.59%	
German	47.27%	
Polish	45.76%	
Indonesian	45.44%	
Hungarian	43.83%	
Fries	41.6%	
Spanish	40.85%	
Norwegian	38.14%	
Portuguese	37.52%	
Dutch	35.22%	
Tongonian	34.68%	
Suid Afrikaans	31.06%	
Hawaiian	19.1%	

1b

Bij Fuzzums gaat men af op de graad van waarschijnlijkheid.

Ook **Fuzzums** www.fuzzums.nl/talenknobbel is een taalherkenner, met het verschil dat de output telkens geëvalueerd wordt volgens waarschijnlijkheid. **Fuzzums** beweert dat tekst 1 statistisch voor 62,38% overeenkomt met Lets, voor 53,73% met Engels, voor 52,8% met Latijn en zelfs voor 35,22% met Nederlands (zie afbeelding 1b). Natuurlijk kijken we alleen naar de meest waarschijnlijke oplossingen, en daar kunnen we zien hoe de tweede taalherkenner grandioos buist. Hij herkent de eerste taal als Lets, maar vergist zich bij taal 2 en 3, die hij respectievelijk Zweeds en opnieuw Lets noemt.
We testen de twee talenherkenners met een instinker, en schotelen ze een zin voor uit de West-Vlamsche Wikipedia: "Wikipedia es e project dat de bedoelinge eit in iedre toale ne vriejn en neutroaln encyclopedie te schrivn". Geloof het of niet, maar **Xerox** zegt zonder verpinken dat het om "Dutch" gaat, en ook **Fuzzums** zet onze moedertaal met 70,21% op kop.

Scenario 2: Russisch ontcijferen

Proef 2. We openen de Moskouse online krant **Azeri.ru** www.azeri.ru en kopiëren een voor ons onverstaanbare Russische tekst waar een neutrale foto van een man bij staat (zie afbeelding 2). Bij Machine Trans-

"Бремя власти"

Андрей Кончаловский представил свой взгляд на жизнь Гейдара Алиева



В московском кинотеатре «Баку» состоялась премьера нового документального фильма про судьбу общенационального лидера Азербайджана Гейдара Алиева. Фильм снимался на протяжении пяти лет в Баку, Нахичевани, Москве, и других городах. Благодаря содействию азербайджанских и российских чиновников в фильме удалось показать редкие документальные кадры и фотографии из личного архива... далее »

2

Geen flauw idee waarover dit gaat.

lators www.humanitas-international.org/newstran/comparisons.htm vinden we alle grote vertaalmachines bij elkaar, zodat we ze gemakkelijk kunnen vergelijken. **Babelfish** http://babelfish.altavista.com, een Altavista-product, verslikt zich in het Russisch. Ook **InterTran** www.tranexp.com:2000/InterTran, dat normaal zelfs Russisch naar Nederlands moet kunnen omzetten, moet passen. Bij **Free Translations** www.freetranslation.com hoort Russisch niet tot de mogelijkheden, en bij **Google Translate** http://translate.google.com en **Dictionary.com** www.dictionary.com evenmin. Alleen **WorldLingo** www.worldlingo.com begint aan de klus. We opteren voor de vertaling Russisch-Nederlands, en dit is het ongekuiste resultaat:

De „last van gezag“

Andrey [Konchalovskiy] stelde zijn vooruitzichten op het leven van Heidar Aliev in de bioskoop van Moskou van Baku voor het de eerste minister van nieuwe documentaire film over het lot van de algemene nationale leider van Azerbaijan Heidar [alieva] plaatsvond. De film werd verwijderd voor een periode van vijf jaar in Baku, Nakhichevan', Aan Moskou, en andere steden. Wegens de hulp van Azerbaijan en Russische ambtenaren in de film was het mogelijk om zeldzame documentaire personeel en foto's van hem persoonlijk te tonen [arkhi]

Bij de **Advanced options** kunnen we instellen hoe speciale lettertekens moeten gelezen worden. Bovendien wordt gevraagd over welk tekstonderwerp het gaat: auto, luchtvaart, chemie, … Naargelang het thema dat je kiest, spreekt **WorldLingo** een ander woordenboek aan.

Het resultaat van deze Russische krachtproef kunnen we moeilijk een behoorlijke vertaling noemen, maar toch kunnen we uit de krampachtige conversie opmaken dat het gaat over een eerder verboden documentaire van Konchalovskiy.

Systran, dat door de zoekmachine **AltaVista** wordt gebruikt onder de naam **BabelFish**, is het oudste commerciële vertaalsysteem. Het dateert zelfs uit de zeventiger jaren. Ook **Dictionary.com** gebruikt **Systran**. **WordLingo** is nog niet zo lang op het web aanwezig, en de vertaalkwaliteit naar het Nederlands is beduidend slechter dan naar het Engels. Dat komt omdat het programma, wanneer je bijvoorbeeld van Russisch naar Nederlands vertaalt, eerst intern een Engelse versie maakt en daarna nog een keertje overschakelt naar het Nederlands. De kwaliteit van **InterTran**-vertalingen is gewoon slecht, maar men kan hier wel Nederlands aan. Met veel goede wil begrijpen we dankzij WordLingo wat er bedoeld wordt, maar het lijkt wel op een Rus die toeristen-Nederlands spreekt.

Scenario 3: Keurig naar het Engels

Over naar proef 3. Hoe doen de online vertalers het wanneer we een stukje tekst naar behoorlijk Engels willen omzetten? We kozen een artikel van ZDNet, geschreven door Jamie Biesemans. Dat begint met:

"Je geslacht heeft een grotere invloed op je IT-vaardigheden dan je opleiding. Dat is de verrassende conclusie van een nieuw Belgisch onderzoek onder studenten..."

Free Translation

Free Text Translator | Free Web Translator

Your FREE Translation *

in a standard way an investigation of sociologists Jef Peeraer and Jef Verhoeven. Peeraer and Verhoeven interrogated 2,200 freshman students. More than 97 per cent of the students uses the computer weekly, the half daily. The pc becomes uses sleepiness for recreation goals, for instance round to chatten or to gamen. Almost ninety per cent start weekly a chatprogramma on. What not per se a bad affair

Have this text professionally translated for only $50.00

Email this translation to yourself or friends...

Translate again:

Dutch to English

Je geslacht heeft een grotere invloed op je IT-vaardigheden dan je opleiding. Dat is de verrassende conclusie van een nieuw Belgisch onderzoek onder studenten. Jongens hebben hun voorsprong op gebied van IT-vaardigheden mogelijk te danken aan games en andere vormen van recreatie op de pc, citeert De Standaard een onderzoek van sociologen Jef Peeraer en Jef Verhoeven.

Other

[Advanced Translation Options]

Human Translation | FREE Translation

3

Wie een professionele vertaling wil, moet 50 dollar betalen.

We gaan de complete vertaling niet weergeven, maar het eerste zinnetje zegt genoeg:

BabelFish: You line has a larger influence on your It-vaardigheden then your training.

WorlLingo: You line has a larger influence on your IT-vaardigheden then your training.

InterTran: Yours tribe has one bigger pull worn yours IT-skill then yours training.

Dictionary.com: geeft een foutmelding

Niemand slaagt er in om de tekst in keurig Engels om te zetten. Wel merken we dat de resultaten nauwkeuriger worden als we de tekst paragraaf per paragraaf in plaats van in zijn geheel laten vertalen.

Scenario 4: Websitevertaling

In onze laatste test gebruiken de websitevertaler van **InterTran**. We stellen die zo in dat we de Engelse website www.zdnet.com vertaald willen zien in het Nederlands (zie afbeelding 4). Het resultaat is hilarisch: 'reviews' wordt 'wapenschouwing', 'file' wordt 'vijl' en 'Microsoft' wordt 'microscoop'.

4

Rechts de vertaalde versie van ZDNet.

Rekenen op een online vertaler is zoiets als vertrouwen op de neef van de kelner in een Maleisisch dorp die beweert "teacher English" te zijn. Met veel goede wil versta je ongeveer de helft van wat er wordt bedoeld. En op vakantie steekt het allemaal niet zo nauw. Maar daarmee is alles gezegd. Tenzij nog dit: don't forget the guide, n'oubliez pas le guide, non dimenticare la guida! ♦